Structural characterization of a protein/RNA complex: human TAP/NXF1 protein/retroviral CTE RNA

Dissertation submitted to obtain the degree of Philosophy Doctor (PhD)
as sponsored jointly by

The European Molecular Biology Laboratory (EMBL)
International PhD Programme

and

The Combined Faculties for the Natural Sciences and for Mathematics
of the Rupertus Carolus University of Hiedelberg, Germany

Presented by
Genaro Pimienta
Guadalajara, Mexico

Heidelberg, 2004

Supervisor: Dr. Elena Conti

Referees

Prof. Dr. Elisa Izaurralde
Prof. Dr. Michael Sattler
Prof. Dr. Klaus Scheffzek
Prof. Dr. Irmgard Sinning

Acknowledgements

I am endlessly grateful to the EMBL PhD Programme, for giving me the opportunity of working on my PhD thesis, within such a unique academic environment. Especially to Elena Conti, who hosted and guided my research work, with much patience.

I wish to thank Katia Zanier, Bernd Simon and Michael Sattler, from the NMR group at the EMBL-Heidelberg, and Dmitry Svergun from the EMBL-Hamburg, who provided support, in a major part of my experimental work. Elisa Izaurralde, provided most of the clones, I have used during this thesis. Judith Ebert introduced me to the project, which was initiated by Erika Liker, who started working with TAP, before I joined the laboratory of Elena Conti.

Many people gravitated on my personal sphere, during the last 4 years. I will abstain myself from outlining a list of names, to avoid omissions and occasional overestimations, and rather keep in tone with the every day cultivation of their friendship, as it has been until today.

To pursue with this PhD thesis, I received a 4-years fellowship from CONACyT, Mexico and a “topping-up” stipend from the EMBL-Heidelberg.

Abstract

Eukaryotic gene expression involves several steps, including transcription, post-transcriptional processing of pre-mRNA transcripts and export of the correctly processed mRNAs from the nucleus to the cytoplasm, where translation takes place. Evidence has accumulated suggesting that these steps are both functionally and physically coupled by protein-protein interactions, involving proteins that bind the mRNA, packing it into ribonucleoprotein particles (mRNPs).

In humans, the protein TAP/NXF1 interacts with mRNA export cargoes and mediates their nuclear export by shuttling through nuclear pore complexes (NPCs). TAP is a multidomain modular protein composed of an N-terminal mRNP-binding region and a C-terminal NPC-binding region. TAP is believed to contact the cellular mRNA indirectly via protein adaptors such as REF/Aly. It can also mediate the nuclear export of exogenous and partially processed viral mRNAs containing a cis-acting Constitutive Transport Element (CTE), which is the case for some types of simian retroviruses. The CTE RNA binds directly to the N-terminal region of TAP and accesses the cellular mRNA export pathway, skipping several upstream events of the mRNA maturation process (e.g. splicing).

The purpose of this work is to characterise the surface and mode of interaction of TAP N-terminal with the CTE RNA and some of its putative cellular adaptors, in particular the mammalian protein REF/Aly. Given this information, we hope to advance in understanding how TAP couples late stages of pre-mRNA processing and mRNA nuclear export. We have used for this purpose, X-ray crystallography and multidimensional NMR spectroscopy, in combination with small angle X-ray scattering experiments.

Table of content

I. Introduction……………………………………………………………………………8

General Overview…………………………………………………………………………..8

Chapter 1. Basics on general nucleocytoplasmic transport…………………………...9

1.1) Nucleocytoplasmic transport…………………………………………………….. 9

1.2) The NPC structure…………………………………….…………………………..9

1.3) Translocation across the NPC……………………………………………………...11

1.4) The nuclear transport receptors…………………………………………………..12

1.5) Karyopherins: the importin β-like family of transport receptors………………...13

1.5.1) Structure of Karyopherins……………………………………………………...16

1.5.2) Ran: the regulator of Karyopherin-mediated nuclear transport…………………...16

1.2) NTF2: the transport receptor of RanGDP……………………………………..18

Chapter 2. Biogenesis and nuclear export of cellular mRNAtranscript…………….19

2.1) Biogenesis of eukaryotic mRNA transcripts……………………………………..19

2.2) TAP/NXF1: the export factor receptor of most cellular mRNA………………....22

2.2.1) Structure of hTAP/NXF1………………………………………………………23

2.3) The REF family of protein adaptors……………………………………………..26

2.3.1) The structure of REF/Aly………………………………………………………28

Chapter3. Biogenesis and nuclear export of retroviral mRNA transcripts…………………………………………………………………………………29

3.1) Retroviruses………………………………………………………………………29

3.2) Complex retroviruses: HIV-1 and the karyopherin mediated export……………….30

3.3) The structure of Rev…………………………………………………………..30

3.4) Simple retroviruses: Simian type D retroviruses and TAP-mediated export……….32

Chapter 4. The questions addressed in this PhD thesis…………….………………...33

1) Main project…………………………………………………………………………...33

2) Side project……………………………………………………………………………33

II. Results and discussion……………………………………………………………35

II.1 Main project……………………………………………………………………..35

Structural characterization of a protein/RNA complex: human TAP/NXF1 protein/retroviral CTE RNA

1) Optimization of an *in vitro* TAP/CTE reconstitution protocol…….……………...36

1.1) RNA minimal construct………………………………………………………..36

1.2) Optimization of the RNA purification protocol……………………………………38

1.3) Protein minimal construct……………………………………………………...40

1.4) Purification of TAP-N to homogeneity……………………………………………..40

1.5) Preparation of spin-labelled samples for NMR experiments………………………..44

1.6) Determination of the molecular weight of TAP-N…………………………………45

1.6.1) Analytical Ultracentrifugation Experiments………………………………………45

1.6.2) Mass Spectrometry……………………………………………………………..45

1.6.3) NMR spectroscopic characterization of TAP-N…………………………………..46

1.7) Investigation of the factors that influence the *in vitro* reconstitution of TAP-N/CTE-B: the role of pH, temperature, salt and protein/RNA molar ratio………………………47

1.7.1) Electrophoretic band shift assays…………………………………………………47

1.7.2) NMR-based titration experiments: monitoring the tendency of TAP-N to oligomerize…………………………………………………………………………47

1.8) Chromatographic characterization of TAP-N, CTE-B and TAP/CTE-B…………...50

1.9) Complex *in vitro* reconstitution……………………………………………………..51

1.10) Stability of the complex in solution………………………………………………..51

Conclusions……………………………………………………………………………53

2) Structural characterization of the TAP-N/CTE-B complex……….……………...54

2.1) Crystallization of TAP-N/CTE-B.…………………………………………………..54

2.2) Multidimensional Nuclear Magnetic Resonance (NMR): the surface on TAP-N that recognizes the CTE-B…………………………………………………………………55

2.2.1) Backbone chemical shift assignments…………………………………………….56

2.2.2) Chemical shift perturbation experiments………………………………………….56

Conclusions……………………………………………………………………………61

2.3) Small angle X-ray scattering (SAXS): the domain arrangement of TAP-N unbound and in complex with the CTE-B RNA…………………………………………………...64

2.4) Overall conclusions: taking together the data obtained from multidimensional NMR and SAXS experiments……………………………………………………………..65

II.2) Side project………………………………………………………………………...66

X-ray crystal structure and ligand-binding surface of REF1-II, a splice variant of the protein REF/Aly

1) Structural characterization of REF1-II………..…………………………………...67

1.1) Sample preparation……………………………………………………………….68

1.1.1) Purification strategy……………………………………………………………..68

1.1.2) Characterization of the REF1-II constructs: stability in solution…………………68

1.2) Crystallographic experiments……………………………………………………..71

1.2.1) Structure determination………………………………………………………...73

Conclusions………………………………………………………………………………74

2) Chemical shift perturbation experiments……………………………………………..79

Conclusions………………………………………………………………………………83

III. Materials and Methods………...………………………………………………..87

General considerations……………………………………………….………………...87

1) Preparation of RNA samples………………………………………………………..88

1.1) CTE RNA constructs ……………………………………………………………88

1.2) Run-off *in vitro* transcription………………………………………………………..88

1.2.1) Large-scale preparation of CTE-encoding pBKS$^+$ plasmids……………………...89

1.2.2) Plasmid linearization……………………………………………………………...89

1.2.3) Run-off *in vitro* transcription reaction……………………………………………90

1.2.4) RNA purification.…………………………………………………………………90

1.2.5) Phenol/chloroform extraction protocol for DNA or RNA, from reaction mixtures…………………………………………………………………………………..91

2) Preparation of protein samples……………………………………………………..93

2.1) Subcloning: reaction conditions and primer oligonucleotides used...………………93

2.1.1) Insert amplification……………………………………………………………..93

2.1.2) Insert purification and ligation into pGEXcs……………………………………...94

2.2) Expression strategy: bacterial strains and plasmids…………………………………...94
2.2.1) Overexpression in LB media……………………………………………………...95
2.2.2) Overexpression in M9 minimal media……………………………………………96
2.2.3) Overexpression of ^{2}H-, ^{15}N-, ^{13}C- TAP-N………………………………………..96
2.3) Purification of recombinant proteins………………………………………………..97
3) Analytical methods…………………………………………………………………..99
3.1) Electrophoresis……………………………………………………………………...99
3.1.1) Denaturing SDS polyacrilamide gel electrophoresis (SDS-PAGE)………………99
3.1.2) Denaturing 8M urea polyacrilamide gel electrophoresis (urea-PAGE)…………100
3.1.3) Native polyacrylamide gel electrophoresis (native-PAGE)……………………..101
3.2) Analytical ultracentrifugation……………………………………………………...102
3.3) Limited proteolysis………………………………………………………………...102
3.4) Protein/RNA binding reaction……………………………………………………..102
4) Biophysical methods………………………………………………………………..103
4.1) NMR measurements………………………………………………………………103
4.2) X-ray data collection………………………………………………………………106
4.3) SAXS data collection………………………………………………………………106
IV. Appendix…..……………………………………………………………………107
A. Sequence alignment of the NXF ortologues in humans……………………………..108
B. Sequence alignment of the REF family of proteins…………………………………109
V. Abbreviations……....………………………………………………….……………..110
VI. References……..…………………………………………………………………..112